

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-102933

(43)Date of publication of application : 16.04.1996

(51)Int.Cl.

H04N 7/173

(21)Application number : 07-103303

(71)Applicant : INTERNATL BUSINESS MACH  
CORP <IBM>

(22)Date of filing : 27.04.1995

(72)Inventor : SHACHNAI HADAS  
WOLF JOEL L  
YU PHILIP S

(30)Priority

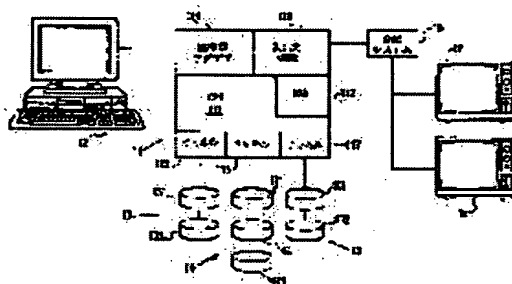
Priority number : 94 241087    Priority date : 11.05.1994    Priority country : US

## (54) BATON PASSING OPTIMIZATION SCHEME FOR LOAD BALANCING/ CONFIGURATION PLANNING IN VIDEO ON DEMAND COMPUTER SYSTEM

(57)Abstract:

PURPOSE: To optimize load balancing by a computer system periodically executing a reallocation function and performing video transfer between disks based on a use's reproduction request to each video.

CONSTITUTION: An operator communicates with a computer through a central control console 12 that is connected to a CPU 111 through an operational console adaptor 114. The CPU 111 is connected to plural direct access storage device DASD strings 13 to 15 through plural channels 115 to 117 respectively. Each DASD string 13 to 15 is connected to one or plural DASDs through a daisy chain, and the number of the DASDs on one string changes a channel number in accordance with a specific introducing system. The DASDs store mass data, and data that is read from a DASD is stored in system memory 12 and read from there to an input-output function 118. A data stream from the function 118 is outputted to a distribution system 16 which supplies it to television sets 17 and 18.



### LEGAL STATUS

[Date of request for examination] 28.10.1997

[Date of sending the examiner's decision of rejection] 04.07.2000

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

BEST AVAILABLE COPY

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-102933

(43) 公開日 平成8年(1996)4月16日

(51) Int.Cl.<sup>6</sup>

H 0 4 N 7/173

識別記号

庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数10 O L (全 15 頁)

(21) 出願番号 特願平7-103303

(22) 出願日 平成7年(1995)4月27日

(31) 優先権主張番号 2 4 1 0 8 7

(32) 優先日 1994年5月11日

(33) 優先権主張国 米国 (US)

(71) 出願人 390009531

インターナショナル・ビジネス・マシーンズ・コーポレーション

INTERNATIONAL BUSINESS MACHINES CORPORATION

アメリカ合衆国10504、ニューヨーク州アーモンク (番地なし)

(72) 発明者 ハダス・シャチナイ

アメリカ合衆国10589 ニューヨーク州ソマーズ クリスタル・ドライブ 46

(74) 代理人 弁理士 合田 潔 (外2名)

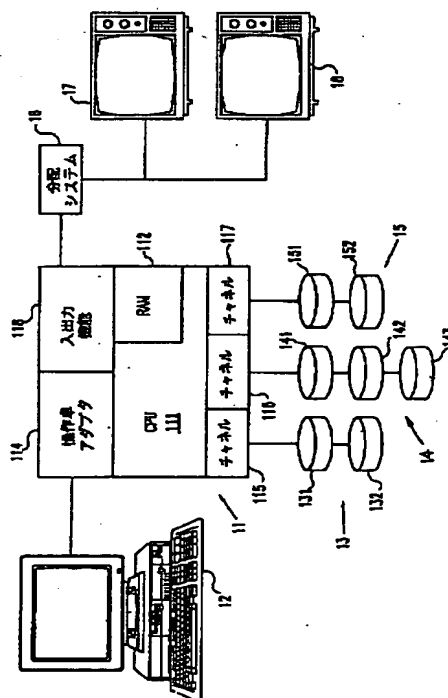
最終頁に続く

(54) 【発明の名称】 ビデオ・オン・デマンド・コンピュータ・システムでの負荷平衡化／構成計画問題のためのパト  
ン渡し最適化方式

(57) 【要約】 (修正有)

【目的】 ディスクに対する負荷を最小化し、平衡化する。

【構成】 ビデオ・オン・デマンド・コンピュータ・システムに、それぞれが複数のビデオを記憶する複数の記憶媒体が含まれる。記憶媒体は、コンピュータ・システムに取り付けられたディスクである。コンピュータ・システムは、データ・ストリームとしてディスクからビデオを読み出すことによってオン・デマンドでビデオを再生して、受け取ったユーザ再生要求に応答してユーザのためにビデオのうちの選択された1つを再生する。コンピュータ・システムは、ディスクのそれぞれについて再生されるビデオの数を監視するようにプログラムされる。コンピュータ・システムによって実行される監視機能に基づいて、コンピュータ・システムは、ディスクのうちの1つから進行中のビデオのコピーを有する別のディスクへ進行中のビデオの再生を転送することによって、負荷平衡化機能を実行する。



1

2

## 【 特許請求の範囲】

【請求項1】 有限の数のビデオを記憶するためのメモリを有するマルチメディア・ファイル・サーバと、それぞれが複数のビデオを記憶する、前記マルチメディア・ファイル・サーバに接続された複数の記憶媒体と、受け取ったユーザ再生要求に応答して、ユーザのためにビデオのうちの選択された1つを再生するための、前記マルチメディア・ファイル・サーバに接続された再生手段とを含む、

前記マルチメディア・ファイル・サーバが、記憶媒体のそれぞれに関して再生中のビデオの数を監視するようプログラムされ、プログラムされた監視動作に基づいて、複数の記憶媒体の負荷を平衡化するために進行中のビデオの再生を記憶媒体のうちの1つから進行中のビデオのコピーを有するもう1つの記憶媒体に転送する中央処理装置を含み、前記中央処理装置が、さらに、ビデオのそれぞれに関するユーザ再生要求に基づいて、負荷平衡化を最適化するために、メモリと記憶媒体との間でビデオを転送することによる前記システム・メモリと記憶媒体との間でのビデオの静的再割当てと、記憶媒体間でビデオを転送することによる記憶媒体間でのビデオの静的再割当てとを周期的に実行するようにプログラムされることを特徴とするビデオ・オン・デマンド・コンピュータ・システム。

【請求項2】 前記中央処理装置が、各シフトが記憶媒体のうちの1つ(ソース)から、シフトされるビデオのコピーを有する別の記憶媒体(ターゲット)へのビデオの再生の転送を引き起こし、ソース記憶媒体が次の負荷シフトのターゲット記憶媒体となる、異なる記憶媒体にまたがる負荷シフトのシーケンスを介して、記憶媒体の任意の対の間で負荷平衡化を実行することを特徴とする、請求項1に記載のコンピュータ・システム。

【請求項3】 前記中央処理装置によって実行される負荷平衡化が、過負荷の記憶媒体と負荷不足の記憶媒体との対の間での負荷平衡化を介する記憶媒体にまたがる負荷の平衡化によるものであることを特徴とする、請求項2に記載のコンピュータ・システム。

【請求項4】 中央処理装置による静的再割当てが、任意の対の記憶媒体の間での負荷シフト・シーケンスの使用可能性を増大させるための記憶媒体へのビデオの割当てと、シーケンスの長さの短縮とによって実行されることを特徴とする、請求項3に記載のコンピュータ・システム。

【請求項5】 システムのコストを最小にするために、所定の数のビデオに関するシステム・メモリおよび記憶媒体の要件を決定するための構成計画手段をさらに含む、請求項1に記載のコンピュータ・システム。

【請求項6】 ビデオに対する需要の予測に従ってシステム・メモリと複数の記憶媒体とにビデオを割り当てるステップであって、前記記憶媒体のそれぞれが複数のビデ

オを記憶し、前記割当てに、単一の記憶媒体上に任意のビデオの1つのコピーだけを有する前記記憶媒体に記憶される各ビデオのコピーの数の決定が含まれる、割当てステップと、

ユーザ再生要求に応答して、ユーザのためにオン・デマンドでビデオのうちの選択された1つを再生するステップと、

記憶媒体のそれぞれについて再生中のビデオの数を監視するステップと、

進行中のビデオの再生を、記憶媒体のうちの1つから、進行中のビデオのコピーを有する別の記憶媒体に転送することによって、前記記憶媒体の負荷を動的に平衡化するステップとを含む、マルチメディア・コンピュータ・システムでビデオ・オン・デマンドを提供する方法。

【請求項7】 ビデオのそれぞれに関するユーザ再生要求に基づいて、負荷平衡化を最適化するために、メモリと記憶媒体との間でビデオを転送することによる前記システム・メモリと記憶媒体との間でのビデオの静的再割当てと、記憶媒体間でビデオを転送することによる記憶媒体間でのビデオの静的再割当てとを周期的に実行するステップをさらに含む、請求項6に記載の方法。

【請求項8】 動的負荷平衡化のステップが、各シフトが記憶媒体のうちの1つ(ソース)から、シフトされるビデオのコピーを有する別の記憶媒体(ターゲット)へのビデオの再生の転送を引き起こし、ソース記憶媒体が次の負荷シフトのターゲット記憶媒体となる、異なる記憶媒体にまたがる負荷シフトのシーケンスを介して実行されることを特徴とする、請求項7に記載の方法。

【請求項9】 動的負荷平衡化のステップが、さらに、過負荷の記憶媒体と負荷不足の記憶媒体との対の間での負荷平衡化によって実行されることを特徴とする、請求項8に記載の方法。

【請求項10】 静的再割当てを実行するステップが、任意の対の記憶媒体の間での負荷シフト・シーケンスの使用可能性を増大させるための記憶媒体へのビデオの割当てと、シーケンスの長さの短縮とによって実行されることを特徴とする、請求項9に記載の方法。

## 【 発明の詳細な説明】

## 【 0 0 0 1 】

【産業上の利用分野】 本発明は、全般的にはマルチメディア・コンピュータ・システム内のファイル・サーバによる記憶装置の管理に関し、具体的には、ムービー・オン・デマンド・コンピュータ・システムと称する場合もあるビデオ・オン・デマンド・コンピュータ・システムでの負荷平衡化/構成計画のための最適化方式に関する。

## 【 0 0 0 2 】

【従来の技術】 ペイ・フォー・ビュー・ムービーは、ケーブル・テレビ・オペレータやホテルが契約者や宿泊客のために提供する、人気が高まりつつあるサービスであ

3

る。ペイ・フォー・ビュー・ムービーは、通常はアナログ形式でビデオ・カセットに記憶される。ビデオ・カセットは、所定の時刻にムービーを再生するようにプログラムされたカルーセルにロードすることができる。セット・トップ・コンバータを有する契約者または宿泊客は、ムービーを選択することができ、そのムービーが、プログラムされた時刻に契約者または宿泊客のテレビ・セットで再生される。契約者または宿泊客は、サービス料金を自動的に請求される。

【 0 0 0 3 】しかし、このサービスは、契約者や宿泊客にとって必ずしも便利とは限らない。というのは、ムービーの選択肢が特定の時間枠にスケジューリングされ、それが契約者や宿泊客のスケジュールに合わない場合があるからである。最近では、契約者や宿泊客が使用可能なムービーのメニューからムービーを選択していつでも自分のテレビで再生できるムービー・オン・デマンド・サービスが提供されつつある。これは、ムービーをデジタル形式で磁気ディスクまたはシステムのランダム・アクセス・メモリ ( RAM ) に記憶することによって可能になった。ディスクに記憶されたムービーは、まずシステム・メモリに読み込まれ、その後、契約者または宿泊客のテレビ・セットに供給されるアナログ・テレビ信号の生成に使用されるデータ・ストリームとして読み出される。契約者または宿泊客のために再生されるムービーがシステム・メモリに静的に記憶される場合、読出しは、ペイ・フォー・ビュー・システムでの動的アナログ記憶によって代表される制約なしに、いつでも行うことができる。

【 0 0 0 4 】ムービー・オン・デマンド・コンピュータ・システムは、多数のムービーの複数のストリームを同時に「再生」できなければならない。これらのムービーは、メモリまたはシステム内の1つまたは複数のディスクに常駐させることができる。現在使用可能な所与のどのムービーについても、期待される需要に関して予測を行うことはできるが、需要は大きく変化する可能性がある。これがシステムの性能を低下させる複数の問題を引き起こす可能性がある。たとえば、最も人気のある3本のムービーをディスクからシステム・メモリにロードして再生することができるが、システム・メモリの制約のために、再生に使用可能な他のムービーはディスクに常駐する。システム・メモリの一部が、契約者または宿泊客がムービーを要求した際にディスクからムービーを「ページング」するために予約される。人気のあるムービーの複数のコピーが、複数のディスクに常駐するが、コピーの数は、そのムービーの現在の人気に応じて変化する。システム内の複数のディスクに対する負荷または需要を平衡化する方法がなければ、所与のディスクからムービーを読み取る要求の量が、サービス可能な量を上回る可能性があり、不可避免的に、契約者または宿泊客に対するサービスの質が低下したり、中断される結果とな

4

る。これは、特に契約者または宿泊客がそのサービスに対して割増料金を支払っている場合には特に、許容できない状況である。

【 0 0 0 5 】

【 発明が解決しようとする課題 】したがって、本発明の目的は、ディスクに対する負荷を最小化し、平衡化するという問題の解決を提供し、したがって、所与のシステム構成に対してより多くの顧客を同時にサービスできるようにすることである。

【 0 0 0 6 】本発明のもう1つの目的は、ムービー・オン・デマンド・システムが最少のコストで再生要件を安全に満たすのに必要なメモリの量またはディスクの量を決定するという関連構成計画問題を解決することである。

【 0 0 0 7 】

【 課題を解決するための手段 】本発明によれば、ディスク負荷の最小化および平衡化の問題は、( 1 ) どのムービーをメモリに常駐させなければならないかと、残りのムービーのコピーをどのディスクに格納するかを静的に決定するステップと、( 2 ) 顧客の需要を再生するためにさまざまなムービーのどのディスク・コピーを使用しなければならないかを動的に決定するステップと、( 3 ) 本明細書でムービー・バトン渡しと称する技法を介してストリームに対する動的調節を実施するステップとによって解決される。この負荷平衡化問題の解決によって、ムービー・オン・デマンド・システムが、所与のメモリとディスクの構成を用いてより多くの顧客負荷をサポートできるようになる。さらに、本発明は、ムービー・オン・デマンド・システムが最少のコストでその再生要件を安全に満足するのに必要なメモリの量およびディスクの量を決定するという関連構成計画問題の解決に使用できる。

【 0 0 0 8 】本発明は、2つの関連するムービー・オン・デマンドの問題を解決する。第1の問題は、入力制約として既存のメモリおよびディスクの構成と、予測されたムービー負荷および実際のムービー負荷とを有する、ディスク負荷平衡化である。この問題の解決は、静的段階と動的段階を有する。静的段階では、ムービーをメモリとディスクに割り当て、動的段階では、ディスクに対する負荷が最少で平衡化された状態でムービーを再生するための方式を提供する。静的段階は、システムの最適リアルタイム動作を保証する動的段階をサポートする。本発明は、本発明の詳細な説明から明らかになるように、「DASDダンス」と称するバトン渡しの処理によって動的段階の負荷平衡化を達成する。静的段階では、( 1 ) DASD対の一方が過負荷であり他方が負荷不足である場合に、バトン渡し手法を使用して、異なるDASDにまたがる負荷シフトのシーケンスを作成することによって過負荷のDASDと負荷不足のDASDの間で負荷を平衡化できる確度を最大化し、( 2 ) 必要な負荷

5

シフトの回数を最少にすることができる形でムービー複製のDASD配置を決定する。負荷不足DASDに記憶されるムービーは、過負荷のDASDに記憶されるムービーと完全に異なる可能性がある。しかし、異なるDASDにまたがるシフトのシーケンスから負荷平衡化を達成でき、この場合、シフトのそれぞれによって、あるDASD(ソース)から、シフトされるムービーのコピーを有する別のDASD(ターゲット)へビデオの実行が転送され、次の負荷シフトではこのソースがターゲットDASDになる。したがって、負荷不足DASDは、最初の負荷シフトのターゲットであり、過負荷DASDは、最後のシフトのソースである。シーケンス内の中間DASDに対する最終負荷は、変化しない。

【0009】第1の問題に対する解決の目標は、固定コストの構成に対して、負荷を処理する能力を最大化することである。第2の関連する問題は、入力制約として予想されたムービー負荷とメモリおよびディスクのコストを有する、構成計画である。この問題に対する解決は、システム・コストを最小にしながらムービー再生要件に合致する構成である。この解決の目標は、固定負荷に対して構成コストを最小化することである。

【0010】

【実施例】ここで図面、具体的には図1を参照すると、本発明を実施することのできるコンピュータ・システムが示されている。このコンピュータ・システムは、たとえばInternational Business Machines (IBM) Corp.社のSystem 390クラスのコンピュータなどのメインフレーム・コンピュータまたは、より新しいIBM Scalable Parallel Systems 9076のうちの1つ、もしくは、本発明のムービー・オン・デマンド機能をサポートするのに十分な入出力機能をサポートできる同様の他のコンピュータ・システムとすることができる。このコンピュータ・システムには、システム・メモリ(RAM)112に接続された中央処理装置(CPU)111を収納するシステム・ユニット11が含まれる。操作員は、操作卓アダプタ114を介してCPU111に接続された中央制御卓12を介してコンピュータと通信する。CPU111は、複数のチャネル115、116および117によって、それぞれ複数の直接アクセス記憶装置(DASD)ストリング13、14および15に接続される。各DASDストリングは、デジタイズ・チェーンで1つまたは複数のDASDを接続されている。チャネルの数と1ストリング上のDASDの数は、特定の導入システムに応じて変化する。図示の例では、DASDストリング13に、DASD131および132が含まれ、DASDストリング14に、DASD141、142および143が含まれ、DASDストリング15に、DASD151および152が含まれる。DASDは、通常は、大量のデータを記憶することができる磁気ディスクである。DASDから読み取られたデータは、システム・メモリ

6

112に記憶され、そこから、入出力機能118へ読み出される。

【0011】図1に示されたものなどのコンピュータ・システムは、トランザクション処理やデータ・ベース管理を含む他の適用業務に使用されてきた。そのような適用業務は、IBM社のMultiple Virtual Storage (MVS) OSやIBM社のAIX OSなどのオペレーティング・システム(OS)によってサポートされる。AIXは、IBM社のUNIXオペレーティング・システムの拡張版である(UNIXはNovell, Inc.社の登録商標である)。そのような適用業務の特徴は、データがDASDから読み取られ、書き込まれることである。これに対して、現在のムービー・オン・デマンド適用業務では、データはDASDから読み取られるのみである。書き込みは、初期構成の際と、静的段階で選択されたディスクにムービーを書き込む時だけに発生する。しかし、データの変更はなく、単なる複製だけである。さらに、トランザクション処理やデータ・ベース管理適用業務は、高速動作によって機能的に改善されるが、本適用業務は、リアルタイム動作を必要とする。すなわち、ある限界以下の速度は許容されない。入出力機能118からのデータ・ストリームは、ビデオ信号を契約者または宿泊客のテレビ・セット17および18に供給する分配システム16に出力される。

【0012】したがって、ムービー・オン・デマンド適用業務は、メインフレーム・コンピュータ上の従来の適用業務と異なる特定の特性を有する。第1に、DASDからのデータ・ストリームは、読取り専用の論理的に定義された単位である。どのデータ・ストリームも、同一とみなすことができる。すなわち、ビット毎秒データ速度は、読み出されるムービーから独立である。さらに、このシステムによって生成されるデータ・ストリームは、予測の際にもリアルタイムでも、基本的に予測可能である。第2に、ムービー・オン・デマンド・システムのリアルタイム・スケジュール要件は、各ディスクがなんらかの並行データ・ストリームの最大数を満足できることを意味し、負荷平衡化がきびしくなることを意味する。しかし、この要件は、複数の異なるDASDに人気のあるムービーの複数のコピーを作成することによって緩和できる。この柔軟性は、複数のDASD上のムービーの複数のコピーを利用して、下で詳細に説明するボタン渡し処理によってDASDに対する負荷または需要を再分配する本発明によってさらに拡張される。以下では、ボタン渡しの処理を「DASDダンス」と称する場合があるが、その理由は本発明の説明を進めるにつれて明らかになる。

【0013】図1に示されたコンピュータ・システムは、典型的なものであり、現在のまたは既知のコンピュータ・システムおよび入出力装置に関して説明されたが、当業者であれば、本発明がこれら特定のコンピュー

タ・システムおよび入出力装置に制限されないことを理解するであろう。明らかに、コンピュータ・ハードウェアが進歩するにつれて、本発明を実施できる他の新しいプラットフォームが現れる。したがって、残りの説明では、コンピュータ・システムを「マルチメディア・サーバ」と呼称し、ムービーを「ビデオ」と呼称し、記憶装置をDASDではなく「ディスク」と呼称する。ここでいう「ディスク」は、幾何学的な円盤であれ他の構成であれ、データを記憶することのできる記憶装置を意味する。これには、たとえばディスク・アレイ、光磁気ディスクおよび他の技術が含まれる。

【0014】本発明には、2つの構成要素がある。第1の構成要素は、負荷平衡化方式であり、第2の構成要素は、構成計画方式である。負荷平衡化方式は、静的段階と動的段階の2段階で進行する。静的段階では、どのビデオをメモリに常駐させ、どのディスクに残りのビデオのコピーを記憶させるかについての良質の割当てを計算する。これらの割当ては、たとえば1日ごとに、増分式に調節される。最大の需要を有する、できる限り多数のビデオを、システム・メモリ112に置く。システム・メモリ112の制約に応じて、これはたとえば、合計100本またはそれ以上のビデオのうちの3本になり、追加のシステム・メモリを、複数の接続されたディスクからのデータのページングに使用できるようになる可能性がある。システム・メモリ112内のビデオは、これ以上の考慮なしにオン・デマンドで再生できる。

【0015】M本の残りのビデオと、合計D台のディスクがあると仮定する(本明細書では、ディスクは、単一の物理ディスクまたはストライプ化されたディスクのグループを意味する)。静的段階の残りの部分で、良質のビデオ/ディスク割当て行列( $a_{ij}$ )を提供するよう努力する。これは、ビデオiがディスクjに置かれる場合には $a_{ij}=1$ 、そうでない場合には $a_{ij}=0$ の、(0, 1)  $M \times D$ 行列である。この割当てで何がよしとされるかを誘導するために、まず、本方式の動的段階を説明する。この段階では、現在の需要ストリームを実際に処理するためにメモリ内にないビデオのどのディスク・コピーを使用すべきかに関するオンライン判断を処理する。

【0016】図2は、本発明の動的段階を呼び出すさまざまなビデオ事象を示す図である。第1の事象は、ブロック21に示されるように、ビデオの再生を求める要求である。この事象は、データ・ストリームを追加するという効果をもたらす。第2の事象は、ブロック22に示されるように、ビデオの再生の完了である。この事象は、あるデータ・ストリームの終了という効果をもたらす。ビデオの起動と停止のほかに、契約者や宿泊客は、家庭用ビデオ・カセット・レコーダ(VCR)を使用する時と同様に、ビデオ再生の一時停止と再開という選択肢が与えられる。ビデオ・オン・デマンドというディジタル技術によって、この機能が可能になる。したがっ

て、契約者または宿泊客がビデオの再生を一時停止させる時には、ブロック23に示されるように、この事象がデータ・ストリームの中断(すなわち、一時的な終了)という効果をもたらす。後程、契約者または宿泊客が一時停止されたビデオの再生を再開する時には、ブロック24に示されるように、この事象がデータ・ストリームの追加という効果をもたらす。これらの事象のそれぞれが、これから詳細に説明する本発明の動的負荷平衡化段階25を呼び出す。

10 【0017】現在ビデオiの $\lambda_i$ 個のストリームが進行中であり、ディスクj上の $\lambda_{ij}$ 個のストリームに分解されていると仮定する。したがって、 $a_{ij}=0$ の時には、必ず

【数1】

$$\lambda_i = \sum_{j=1}^D \lambda_{ij}$$

【0018】かつ $\lambda_{ij}=0$ である。各ディスクjには、最大数 $L_j$ の並行ストリーム(すなわち、再生中のビデオ)が関連する。この数は、ディスクの性能特性に依存し、必要な固定デッドライン以内にビデオを読み取るというリアルタイム・スケジューリング問題の達成が保証されるように選択される。ディスク上の負荷を平衡化するために、ディスクjごとのペナルティ関数 $F_j$ を、集合{0, ...,  $L_j$ }に対する凸増加関数として定義されるものと仮定する。負荷は、制約(1)  $x_{ij} \in \{0, \dots, L_j\}$ 、(2)

【数2】

$$\sum_{j=1}^D x_{ij} = \lambda_i$$

30 【0019】および(3)  $a_{ij}=0$ ならば $x_{ij}=0$ 、の下で関数

【数3】

$$\sum_{j=1}^D F_j \left( \sum_{i=1}^M x_{ij} \right)$$

【0020】が最小化される時に、所与の現在負荷およびビデオ/ディスク割当てに関してできる限り平衡化される。この問題は、いわゆるクラス制約付き資源配分問題であり、たとえばタンタウィ(A. Tantawi)、タウズレイ(D. Towsley)およびウルフ(J. Wolf)共著、"Optimal Allocation of Multiple Class Resources in Computer Systems," ACM Sigmetrics Conference, Santa Fe NM, 1988に記載のグラフ理論最適化アルゴリズムなどを使用して正確かつ効率的に解くことができる。最適解の場合、

【数4】

$$x_{ij} = \sum_{i=1}^M x_{ij}$$

【0021】が、ディスクjに対する所望の負荷を表す。したがって、値

50 【数5】

$$x_j - \sum_{i=1}^M \lambda_{ij}$$

【0022】の昇順でディスクをインデクシングし直すと、負荷が最大のものから最小のものへの順でディスクがならぶ。理想的には、これらの値のすべてが0に近くなければならない。表記を固定するために、最初のD<sub>1</sub>台のディスクが過負荷であり、最後のD<sub>2</sub>台のディスクが負荷不足であると仮定する。

【0023】この説明で使用する表記法を要約する。

- ・ 非メモリ常駐ビデオ  $i = 1, \dots, M$
- ・ ディスク  $j = 1, \dots, D$
- ・ ビデオ/ディスク割当て行列  $(a_{ij})$
- ・ ディスクごとの最大ストリーム負荷  $L_j$
- ・ ディスク・ペナルティ関数  $F_j(\cdot)$  (凸関数である)
- ・ 現在の負荷

【数6】

$$\lambda_i = \sum_{j=1}^D \lambda_{ij}$$

【0024】ただし、 $\lambda_{ij}$ は、ディスク  $j$  上でのビデオ  $i$  の「再生」の数であり、 $\lambda_i$ は、ビデオ  $i$  に対する需

【0025】ここで図3を参照すると、本発明による動的段階の処理が示されている。第1ステップ31で、上で定義した  $\lambda_i$  と  $\{\lambda_{ij}\}$  を調節する。次に、上記のタ

【0026】ストリーム需要は、図2に示されるように、契約者または宿泊客が新しいビデオをスタートさせるか現在一時停止しているビデオを再開する時に、1つ増分される。同様に、ストリーム需要は、契約者または宿泊客がビデオを終了させるか現在再生中のビデオを一時停止させる時に、1つ減分される。通常、需要増加の処理は、貪欲 (greedy) アルゴリズムすなわち、大域的な視野を無視して局所的な最適解をもたらすアルゴリズムを使用することによって達成できる。言い換えると、ビデオ  $i$  の新しいストリームを追加しなければならない場合、

【数7】

$$a_{ij} = 1$$

【0027】を満足し、最初の差

【数8】

$$F_j\left(\left(\sum_{i=1}^M \lambda_{ij}\right) + 1\right) - F_j\left(\sum_{i=1}^M \lambda_{ij}\right)$$

【0028】が最小のディスク  $j$  が選択される。しかし、この手法は周期的に品位が低下する。値

【数9】

$$\left| x_j - \sum_{i=1}^M \lambda_{ij} \right|$$

【0029】と0との差が、合理的な基準によって固定された閾値  $T$  を超える場合、動的段階アルゴリズムが開始される。たとえば、この基準は、二乗和

【数10】

$$B = \sum_{j=1}^D \left( x_j - \sum_{i=1}^M \lambda_{ij} \right)^2$$

【0030】として計算することができる。開始されるアルゴリズムも、グラフ理論によるものであり、下記のように定義される有向グラフ  $G$  を常に維持する。

- ・ ノードはディスクである
- ・ 別個のノードの対  $j_1$  および  $j_2$  のそれぞれについて、

【数11】

$$a_{i,j_1} = a_{i,j_2} = 1$$

【0031】すなわち、ディスク  $j_1$  および  $j_2$  の両方に特定のビデオ  $i$  が存在し、

(2)

【数12】

$$\lambda_{i,j_1} > 0$$

【0032】すなわち、ビデオ  $i$  がディスク  $j_1$  上で再生中であり、

(3)

【数13】

$$\sum_{i=1}^M \lambda_{i,j_2} < L_{j_2}$$

【0033】すなわち、ディスク  $j_2$  が過剰能力を有するビデオ  $i$  が少なくとも1つ存在するならば、 $j_1$  から  $j_2$  への有向弧が存在する。

【0034】図3に戻って、ステップ33で、負荷の昇順でディスクをインデクシングし、不良値  $B$  を計算し、グラフ  $G$  と量  $D_1$  および  $D_2$  を更新する。その後、ステップ34で、 $B$  が  $T$  を超えるかどうかに関する判断を行う。そうでない場合、処理を停止する。そうである場合、ステップ35で  $j_1$  を1にセットし、ステップ36で  $j_2$  を  $D$  にセットする。その後、ステップ37で、 $j_1$  から  $j_2$  への最短経路が存在するかどうかに関する判断を行う。 $j_1$  から  $j_2$  への有向弧の存在は、受取り側ディスクの負荷容量を超えずにあるディスクから別のディスクへあるビデオの単一のストリームをオンラインで転送できる可能性を意味する。 $j_1$  から  $j_2$  への最短経路が存在する場合、ステップ38で、下で説明し図4に示すバトン渡し方式を呼び出し、ステップ33にループ・バックする。 $j_1$  から  $j_2$  への最短経路が存在しない場合、ステップ39で  $j_2$  を減分し、ステップ40で、 $j_2$  が  $D - D_2 + 1$  以上であるかどうかに関する判断を行う。そうである場合、ステップ37にループ・バックする。そうでない場合、ステップ41で  $j_1$  を増分する。 $j_1$  を増分

11

した後に、ステップ42で、 $j_1$ が $D_1$ 以下であるかどうかに関する判断を行う。そうである場合、ステップ36にループ・バックする。そうでない場合、処理を終了する(不成功)。

【0035】図4は、図3のステップ38で呼び出されるボタン渡し処理の流れ図である。ディスク $j_1$ からディスク $j_2$ への最短経路が、1によってインデクシング(索引付け)される $k$ 個のノードからなり、最初のノードがディスク $j_1$ に対応し、 $k$ 番目のノードがディスク $j_2$ に対応すると仮定する。最初のステップ44で、1を $k-1$ にセットする。次に、ステップ45で、ノード1に対応するディスクからノード1+1に対応するディスクへの、共通ビデオ $V_1$ の同期式ボタン渡しを行う。この共通ビデオは、 $G$ の定義によって存在しなければならない。次に、ステップ46で、1を1つ減分する。その後、ステップ47で、1が1以上であるかどうかに関する判断を行う。そうである場合、ステップ45にループ・バックする。そうでない場合、この処理を停止し、図3に示された主プログラムへリターンする。

【0036】弧が1つだけの場合、ボタン渡しは、ある種の同期化プリミティブを介して達成できる。1を超え\*

Procedure SUB:

Do for  $j_1=1$  to  $D_1$

Do for  $j_2=D$  to  $D-D_2+1$  by -1

If  $j_1$ から $j_2$ への最短経路が $G$ に存在する  
then ボタン渡しを使用して  
この経路に沿ってビデオ  
を転送し、(1)を返す

End

End

Return (0)

End SUB

この動的方式は、ビデオを動的に転送しないで可能になるはずの度合より大きい度合まで負荷を平衡化するという効果をもたらす。

【0038】まとめると、本発明による動的段階の方式では、最適化アルゴリズムを介して最適ディスク負荷を決定し、最も負荷の高いものから最も負荷の低いものへの順で(ここでの表記法によれば、最初の $D_1$ 台のディスクが過負荷であり、最後の $D_2$ 台のディスクが負荷不足である)値

【数14】

$$\left| x_j - \sum_{i=1}^M \lambda_{ij} \right|$$

【0039】の昇順でディスクをインデクシングし直す。動的段階では、不良の尺度

【数15】

12

\*る長さの経路の場合、ボタン渡しは逆の順序で、すなわち、最後の弧から最初の弧へと実行しなければならない。これによって、一時的なものも含めてどのディスクでも過負荷がなくなる。全体としての効果は、1つのストリームを経路内の最後のディスクに追加し、1つのストリームを経路内の最初のディスクから取り除くということであり、他のどのディスクにも全く影響がない。この処理を、「逆方向ダンス」と説明することもできる。

【0037】図3および図4に示された処理は、メイン・ルーチンと1つのサブルーチンを有する。メイン・ルーチンは次のとおりである。

Procedure: MAIN

Do while B>T

Call SUB

If SUBが(0)を返す then stop

インデックスB, G, D1, D2に対する簿記を実行する

End

End MAIN

サブルーチンは次のとおりである。

$$B = \sum_{j=1}^D \left( x_j - \sum_{i=1}^M \lambda_{ij} \right)^2$$

【0040】と不良閾値 $T$ を仮定し、不良閾値を超えた時に、改善ルーチンを使用して負荷を調整する。

【0041】動的段階での最適ディスク負荷の質は、ビデオ/ディスク割当てに依存する。図5に、本発明の静的段階の処理を示す。最初のステップ51で、予測された需要の降順でビデオを並べる。静的段階の場合、需要が最も多いビデオがシステム・メモリから再生され、これらのビデオは、ステップ52でシステム・メモリに食欲に割り当てられる。残りの $M$ 本のビデオは、予測された再生需要 $\lambda_i$ の降順でインデクシングされる。したがって、 $\lambda_1 \geq \dots \geq \lambda_M$ である。次に、ステップ53で、そのようなビデオ $i$ のそれぞれについて、必要なコピーの数 $A_i$ を計算する。このステップを行うには、2つの可能な代替案がある。1つの可能性は、 $A$ がマルチメディア・サーバ全体で許容される最大のビデオ・ディスク・



13

コピーの数を表すものとし、問題

【数16】

$$\alpha = \max \left\{ \alpha \sum_{i=1}^M [\alpha \lambda_i] \leq A \right\}$$

【0042】を解くことによって、 $A_i$  がおおまかに  $\lambda_i$  に比例するようにするというものである。この解は、たとえばプレス (W. Press)、フラネリ (B. Flanner)、チューコルスキ (S. Teukolsky) およびベッターリング (W. Vetterling) 共著、Numerical Recipes, Cambridge University Press に記載のものなど、括弧積二分法 (bracket and bisection algorithm) を介して得ることができる。その後、ビデオ  $i$  のそれぞれについて、

$$A_i = [\alpha \lambda_i]$$

【0043】であるものとする。もう1つの可能性は、この問題を公平配分の1つとして扱うことである。配分のための方式は、たとえばイバラキ (T. Ibaraki) およびカトウ (N. Katoh) 著 "Resource Allocation Problems", MIT Press で扱われている。この文献で扱われている適切な方式の1つは、ウェブスター (Webster) によるものである。

【0044】静的段階には2つの可能なモードがある。モード1は、「初めから」であり、まだビデオが一本もディスクに置かれていない新システムを構成するのに使用される。モード2は、「増分式」であり、おおむね毎日の基準で、改訂された予測のためにディスクへの既存のビデオの割当てを調節するのに使用される。この2つの手法の目的は、下記に従って定義される非有向グラフ  $H$  の連結度を高めることである。

- ・ノードはディスクである
- ・別個のノードの対  $i$  および  $j$  のそれぞれについて、(1)

【数18】

$$a_{i,j_1} = a_{i,j_2} = 1$$

【0045】であるビデオ  $i$  が少なくとも1つ存在するならば、 $j_1$  と  $j_2$  の間に弧が存在する。(この条件は、有向グラフ  $G$  の定義の条件(1)を模倣したものである。これは、よい動的段階方式が、通常は条件(1)が満たされる時には必ず  $G$  の定義の条件(2)および(3)が満たされることを保証するので、 $H$  が  $G$  の有効な代理として働くという考えである)

【0046】グラフ  $H$  の「直径」を最小化することによって連結度を高める(グラフの直径とは、ノード対の間の距離の最大値である)。モード1では、いわゆる「クリーク木」を生成することによってこれを達成する。この木のノードのそれぞれは、それぞれが共通のビデオのコピーを有するディスクの集合からなる。これによって、共通ビデオに基づく1ノード内の任意の2台のディスクの間での負荷シフトが可能になる。したがって、ク

14

リーク木という用語を使用する。この木は、処理中にノードにディスクを割り当て、ディスクにビデオを割り当てることによって作成する必要がある。この木は、アクセス頻度の予測値の降順でビデオを検討することによって作成される。最も人気の高いビデオのコピーを、まず異なるディスクに割り当て、これらのディスクを、クリーク木の根ノードに写像する(各ディスクには、ビデオの1つのコピーを格納することができる)。次に、(需要に関して) 次に「人気の高い」複数のコピーを有するビデオを検討する。やはり、異なるコピーを異なるディスクに割り当てるが、コピーのうちの1つは、クリーク木に既に存在するノードのディスクに割当てて必要がある。これによってクリーク木への連結度がもたらされるので、これを連結コピーと称する。負荷を平衡化するための負荷シフトまたはバトン渡しの回数を最少にするために、新しいノードは、できる限り根の近くに付加される。これは、根に最も近く、追加容量があるディスクを有するノードに「連結」コピーを割り当てることによって達成される。この処理は、クリーク木に含まれないままになっているディスクがある限り繰り返される。葉ノードは、複製を有する人気の低いビデオを表す。木作成の最終段階の間に、ビデオが、割り当てられないままになっているディスクより多数のコピー(連結コピーを除く)を有する場合がある。割り当てられないままのディスクの台数を越えるコピーの割当ては、次のステップで対処する。すべてのディスクをクリーク木に含めた後に、ビデオの残りのコピーを、ディスクに食欲に割り当てる。すなわち、(i) 新しいコピーの対のそれぞれについて、使用可能な空間を有するディスクの集合から最大の距離を有するディスク対を選択し、(ii) ビデオの単一のコピーを、直径が減るようにディスクに置く(そのビデオの追加のコピーがある場合、選択されたディスクが、そのビデオの他のコピーを収容するディスクのクリークの要素になることに留意されたい。残りのビデオは、 $A_i = 1$  を満足し、 $H$  の直径に影響しない)。以下では、まずクリーク木を作り、その後グラフの直径を減らすために食欲方式を使用することによる、グラフ  $H$  の構成を示す。

【0047】例: ディスクの台数  $D=11$  で  $L_j=2$ ,  $j=1, \dots, D$  であり、ビデオの数  $M=11$  であり、各ビデオのコピーの数が下の表で与えられるシステムを仮定する。

【表1】

15

| ビデオID | コピー |
|-------|-----|
| A     | 4   |
| B     | 3   |
| C     | 3   |
| D     | 3   |
| E     | 2   |
| F     | 2   |
| G     | 2   |
| H     | 2   |
| I     | 2   |
| J     | 1   |
| K     | 1   |

【 0 0 4 8 】 初期段階( すなわちクリーク木の構築) の後のディスクへのビデオ・コピーの割当ては、次の行列によって与えられる。

【 数1 9 】

$$a_{ij} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

【 0 0 4 9 】 食欲段階の後に得られるグラフHを、図6に示す。上の表と図6を参照すると、ビデオAには4つのコピーがあり、そのそれぞれが、ディスク6 1 1、6 1 2、6 1 3および6 1 4に割り当てられている。この4台のディスクのグループが、クリーク木の根6 1である。次に人気の高いビデオは、3つのコピーを有するビデオBである。この3つのうちの1つは、ディスク6 1 1に割り当てられ、このビデオの連結コピーとして働く。ビデオBの他の2つのコピーは、それぞれディスク6 2 1および6 2 2に割り当てられる。この2台のディスクが、クリーク木の枝6 2を形成する。ビデオCも3つのコピーを有し、そのうちの1つがディスク6 1 2に割り当てられて、ビデオCの他の2つのコピーが割り当てられるディスク6 3 1および6 3 2を有する枝6 3への連結コピーとして働く。同様に、ビデオDには3つのコピーがあり、そのうちの1つが連結コピーとしてディスク6 1 3に割り当てられ、他の2つは、枝6 4のディスク6 4 1および6 4 2に割り当てられる。ビデオEには2つのコピーがあり、一方はディスク6 1 4に割り当てられ、他方はディスク6 5に割り当てられる。この場合、ディスク6 5は、クリーク木の葉を表す。

【 0 0 5 0 】 割当て処理のこの時点で、ディスク6 1 1、6 1 2、6 1 3および6 1 4の容量に達し、その結果、ビデオの次のグループの連結コピーをクリーク木の

16

根6 1に割り当てることはもはや不可能になっている。したがって、ビデオFのコピーの一方は、ディスク6 2 1に割り当てられ、もう一方は、ディスク6 6に割り当てられる。同様に、ビデオGのコピーの一方は、ディスク6 2 2に割り当てられ、もう一方は、ディスク6 7に割り当てられる。ディスク6 6および6 7は、ディスク6 5と同様に、クリーク木の葉である。次に、ビデオHの2つのコピーを、ディスク6 3 1および6 6に割り当て、ビデオIの2つのコピーを、ディスク6 4 1および6 7に割り当てる。これによって、それぞれ1つのコピーを有するビデオJおよびKが残される。ビデオJのコピーは、ディスク6 3 2に割り当てられ、ビデオKのコピーは、ディスク6 4 2に割り当てられる。

【 0 0 5 1 】 前述は、ディスクへのビデオの初期割当ての1例にすぎない。明らかに、ビデオの数、各ビデオのコピーの数、ディスクの台数および容量に応じて、ディスクへのビデオの割当てが変化する。この例のポイントは、バトン渡し技法を使用して、共通するビデオのコピーを有しない可能性がある2台のディスクの間で負荷を平衡化することができる状況を示すことにある。たとえば、ディスク6 1 1が過負荷であり、ディスク6 6が負荷不足であると仮定する。この2台のディスクには共通のビデオのコピーが存在しないことに留意されたい。しかし、ディスク6 2 1は、ディスク6 1 1との間にもディスク6 6との間にも共通のコピーを有する。この場合のバトン渡しは、ビデオFに対する需要をディスク6 2 1からディスク6 6に転送し、その後、ビデオBに対する需要をディスク6 1 1からディスク6 2 1に転送することである。この「逆方向ダンス」によって、一時的なものも含めて中間のディスク6 2 1に対する過負荷状態が生じなくなる。全体としての効果は、ディスク6 6のビデオ・ストリーム出力が1つ増加し、ディスク6 1 1のビデオ・ストリームが1つ減少することであり、ディスク6 2 1のビデオ・ストリーム出力には何の影響もない。

【 0 0 5 2 】 モード2では、一部のビデオのディスク・コピーが減り、一部のビデオのディスク・コピーが増え、一部のビデオ・コピーはディスクからディスクへ移動される可能性がある。ディスク上の新しいビデオ・コピーの総数を追跡し、ある固定された閾値Uがその上限になるようにする。

【 0 0 5 3 】 この手順には3つの段階があり、下記の順序で達成される。段階1では、正味減少ビデオから、減少の大きいものから小さいものへの順でディスク・コピーを食欲に除去する。段階2では、正味増加ビデオから、増加の大きいものから小さいものへの順で、ディスク記憶容量の制約の下でディスク・コピーを食欲に追加する。その後、新しいビデオ・コピーの数を計算する。この数がU以上の場合、処理を停止する。そうでない場合、段階3すなわち、ビデオ・コピーの集合全体に対す

10

20

30

40

50

17

る近隣エスケープ・ヒューリスティックを実行する。使用される近隣エスケープ・ヒューリスティックは、ウルフ(J. Wolf)著、"The Placement Optimization Program", ACM Sigmetrics Conference, Berkeley CA, 1989に記載されている。これには、ディスク記憶容量の制約付きでの、単一移動(距離1)、交換(距離2)などが含まれる。(1) Hの直径のこれ以上の改善が不可能であるか、(2) 移動の数がUを超えるのいずれかの場合には、処理を停止する。後者の場合、最後の移動判断を打ち切る。Hの直径を最小化すると、動的段階方式の柔軟性が増すので、動的段階方式がより効率的になる。複数のコピーを有する適度な本数のビデオであっても、Hの良好な連結度が保証される。

【0054】図5に戻って、ステップ54で、 $A_i^{old}$ が $A_i$ より大きい時には必ず、食欲(GREEDY)アルゴリズムを使用してディスク・コピーを除去する。ここで、 $A_i^{old}$ は、ビデオiの必要なディスク・コピーの古い数であり、 $A_i$ は、新しい数を表す。その後、ステップ55で、 $A_i$ 未満の $A_i^{old}$ のそれぞれについて、 $A_i - A_i^{old}$ の降順でビデオを並べ直す。ステップ56で、 $A_i^{old}$ が $A_i$ 未満の時には必ず、順番にビデオのディスク・コピーを追加する。これは、食欲アルゴリズムを介して実行される。その後、 $\mu$ を

【数20】

$$\sum_{i=1}^M (A_i - A_i^{old})$$

【0055】にセットする(関数+は、 $x \geq 0$ ならば $x^+ = x$ 、 $x \leq 0$ ならば $x^+ = 0$ として定義される)。ステップ57で、 $\mu$ がU以上であるかどうかに関する判断を行う。そうである場合には、処理を停止し、そうでない場合には、ステップ58で、上記のウルフに記載の近隣エスケープ・ヒューリスティックを使用して新しいビデオ・コピーを追加するかどうかに関する判断を行う。そうでない場合には、処理を停止し、そうである場合には、ステップ59でビデオをコピーし、その後、ステップ60で $\mu$ を更新してから、ステップ57にループ・バックする。

【0056】要約すると、静的段階モードは、「初めから」と「増分式」である。「初めから」モードは、新システムを構成する時に限って使用される。予測を使用して、メモリ常駐ビデオと残りのビデオのディスク・コピーの数とを決定する。その手法は、食欲戦略を使用し、ディスク容量に対する制約を考慮に入れてクリーク木を作成することである。「増分式」モードは、通常の日次動作に使用される。「増分式」モードでは、改訂された予測を使用して、メモリ常駐ビデオを更新し、残りのビデオのディスク・コピーの数を調節する。食欲戦略と近隣エスケープ・ヒューリスティックを使用する「増分式」モードには、3部分の手法がある。正味減少ビデオについては、ビデオのどのコピーを消去するかに関して

18

食欲な判断を行う。正味増加ビデオについては、ディスク記憶容量の制約を考慮に入れて、どこに新しいコピーを置くかに関して食欲な判断を行う。近隣ヒューリスティックを使用して、ディスク移動の回数とディスク記憶容量に関する制約を考慮に入れて、すべてのビデオ・コピーを移動する。

【0057】構成計画問題は、本質的に負荷平衡問題と重なるものである。これは、(1) ビデオ負荷シミュレーション・テストの組を開発し、(2) これらのテストに合格するメモリ構成およびディスク構成を探して探索空間を調べ、(3) (2)を満足し最小のコストを有する構成を選択することによって、同一の基本技法を使用して達成できる。

【0058】単一の好ましい実施例に関して本発明を説明してきたが、当業者であれば、請求の範囲の趣旨と範囲の中で本発明に変更を加えて実施できることを諒解するであろう。

【0059】まとめとして、本発明の構成に関して以下の事項を開示する。

【0060】(1) 有限の数のビデオを記憶するためのメモリを有するマルチメディア・ファイル・サーバと、それぞれが複数のビデオを記憶する、前記マルチメディア・ファイル・サーバに接続された複数の記憶媒体と、受け取ったユーザ再生要求に応答して、ユーザのためにビデオのうちの選択された1つを再生するための、前記マルチメディア・ファイル・サーバに接続された再生手段とを含み、前記マルチメディア・ファイル・サーバが、記憶媒体のそれぞれに関して再生中のビデオの数を監視するようプログラムされ、プログラムされた監視動作に基づいて、複数の記憶媒体の負荷を平衡化するために進行中のビデオの再生を記憶媒体のうちの1つから進行中のビデオのコピーを有するもう1つの記憶媒体に転送する中央処理装置を含み、前記中央処理装置が、さらに、ビデオのそれぞれに関するユーザ再生要求に基づいて、負荷平衡化を最適化するために、メモリと記憶媒体との間でビデオを転送することによる前記システム・メモリと記憶媒体との間でのビデオの静的再割当てと、記憶媒体間でビデオを転送することによる記憶媒体間でのビデオの静的再割当てとを周期的に実行するようにプログラムされることを特徴とするビデオ・オン・デマンド・コンピュータ・システム。

(2) 前記中央処理装置が、各シフトが記憶媒体のうちの1つ(ソース)から、シフトされるビデオのコピーを有する別の記憶媒体(ターゲット)へのビデオの再生の転送を引き起こし、ソース記憶媒体が次の負荷シフトのターゲット記憶媒体となる、異なる記憶媒体にまたがる負荷シフトのシーケンスを介して、記憶媒体の任意の間で負荷平衡化を実行することを特徴とする、上記(1)に記載のコンピュータ・システム。

(3) 前記中央処理装置によって実行される負荷平衡化

が、過負荷の記憶媒体と負荷不足の記憶媒体との対の間での負荷平衡化を介する記憶媒体にまたがる負荷の平衡化によるものであることを特徴とする、上記(2)に記載のコンピュータ・システム。

(4) 中央処理装置による静的再割当てが、任意の対の記憶媒体の間での負荷シフト・シーケンスの使用可能性を増大させるための記憶媒体へのビデオの割当てと、シーケンスの長さの短縮とによって実行されることを特徴とする、上記(3)に記載のコンピュータ・システム。

(5) システムのコストを最小にするために、所定の数のビデオに関するシステム・メモリおよび記憶媒体の要件を決定するための構成計画手段をさらに含む、上記(1)に記載のコンピュータ・システム。

(6) ビデオに対する需要の予測に従ってシステム・メモリと複数の記憶媒体とにビデオを割り当てるステップであって、前記記憶媒体のそれぞれが複数のビデオを記憶し、前記割当てに、単一の記憶媒体上に任意のビデオの1つのコピーだけを有する前記記憶媒体に記憶される各ビデオのコピーの数の決定が含まれる、割当てステップと、ユーザ再生要求にตอบสนองして、ユーザのためにオン・デマンドでビデオのうちの選択された1つを再生するステップと、記憶媒体のそれぞれについて再生中のビデオの数を監視するステップと、進行中のビデオの再生を、記憶媒体のうちの1つから、進行中のビデオのコピーを有する別の記憶媒体に転送することによって、前記記憶媒体の負荷を動的に平衡化するステップとを含む、マルチメディア・コンピュータ・システムでビデオ・オン・デマンドを提供する方法。

(7) ビデオのそれぞれに関するユーザ再生要求に基づいて、負荷平衡化を最適化するために、メモリと記憶媒体との間でビデオを転送することによる前記システム・メモリと記憶媒体との間でのビデオの静的再割当てと、記憶媒体間でビデオを転送することによる記憶媒体間でのビデオの静的再割当てとを周期的に実行するステップをさらに含む、上記(6)に記載の方法。

(8) 動的負荷平衡化のステップが、各シフトが記憶媒体のうちの1つ(ソース)から、シフトされるビデオのコピーを有する別の記憶媒体(ターゲット)へのビデオの再生の転送を引き起こし、ソース記憶媒体が次の負荷シフトのターゲット記憶媒体となる、異なる記憶媒体にまたがる負荷シフトのシーケンスを介して実行されることを特徴とする、上記(7)に記載の方法。

(9) 動的負荷平衡化のステップが、さらに、過負荷の記憶媒体と負荷不足の記憶媒体との対の間での負荷平衡化によって実行されることを特徴とする、上記(8)に記載の方法。

(10) 静的再割当てを実行するステップが、任意の対の記憶媒体の間での負荷シフト・シーケンスの使用可能性を増大させるための記憶媒体へのビデオの割当てと、シーケンスの長さの短縮とによって実行されることを特

徴とする、上記(9)に記載の方法。

【 0 0 6 1 】

【 発明の効果】コンピュータ・システムは、再割当て機能を周期的に実行して、ビデオのそれぞれに対するユーザ再生要求に基づいて負荷平衡化を最適化するためにディスク間でビデオを転送する。コンピュータ・システムによって実行される負荷平衡化には、静的段階と動的段階の2段階がある。静的段階では、ビデオをメモリとディスクに割り当て、動的段階では、ディスクに対する負荷が最小で平衡した状態でビデオを再生するための方式を提供する。静的段階は、システムの最適リアルタイム動作を保証する動的段階をサポートする。動的段階の負荷平衡化は、バトン渡しの処理によって達成される。

【 図面の簡単な説明】

【 図1 】本発明が実施されるコンピュータ・システムの機能ブロック図である。

【 図2 】本発明の動的段階とビデオ事象の間の相互作用の機能ブロック図である。

【 図3 】本発明の動的段階の論理の流れ図である。

【 図4 】本発明の動的段階のバトン渡し構成要素の論理の流れ図である。

【 図5 】本発明の静的段階の「増分」方式の論理の流れ図である。

【 図6 】システム内のディスクへのムービーの割当ての例を示す機能ブロック図である。

【 符号の説明】

- 1 1 システム・ユニット
- 1 2 中央制御卓
- 1 3 直接アクセス記憶装置(DASD)ストリング
- 1 4 直接アクセス記憶装置(DASD)ストリング
- 1 5 直接アクセス記憶装置(DASD)ストリング
- 1 6 分配システム
- 1 7 テレビ・セット
- 1 8 テレビ・セット
- 6 1 根
- 6 2 枝
- 6 3 枝
- 6 4 枝
- 6 5 ディスク
- 6 6 ディスク
- 6 7 ディスク
- 1 1 1 中央処理装置(CPU)
- 1 1 2 システム・メモリ(RAM)
- 1 1 4 操作卓アダプタ
- 1 1 5 チャネル
- 1 1 6 チャネル
- 1 1 7 チャネル
- 1 1 8 入出力機能
- 1 3 1 DASD
- 1 3 2 DASD

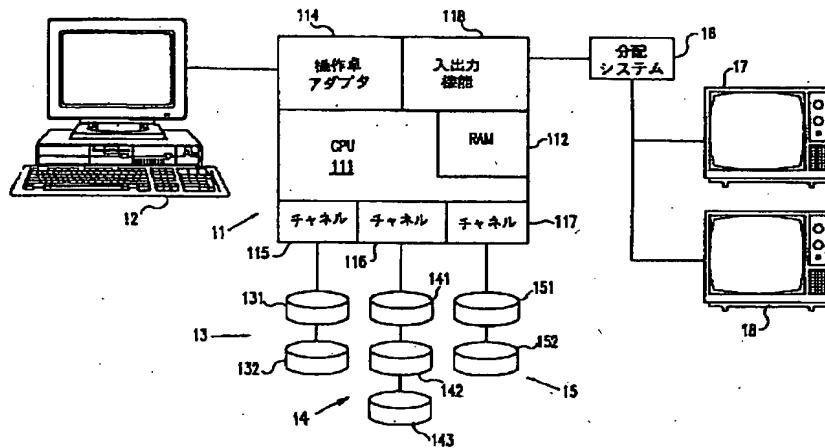
21

22

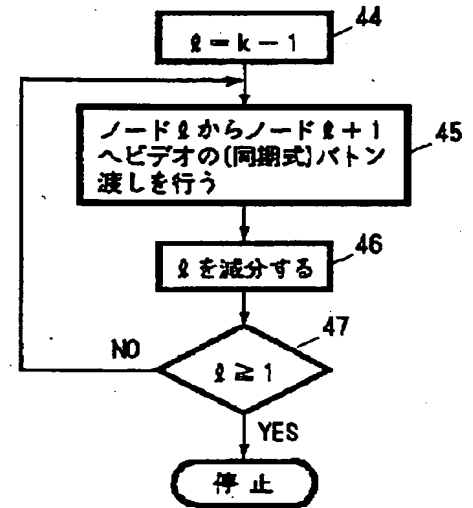
141 DASD  
142 DASD  
143 DASD

151 DASD  
152 DASD

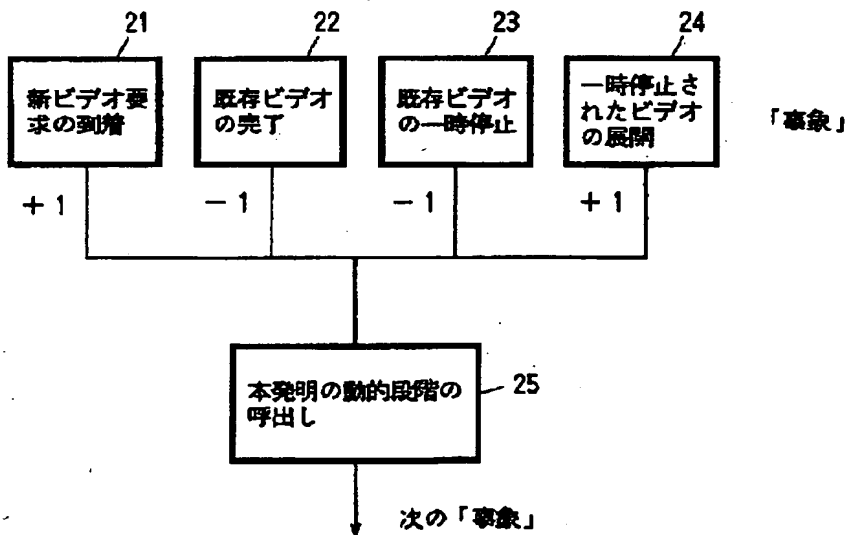
【 図1 】



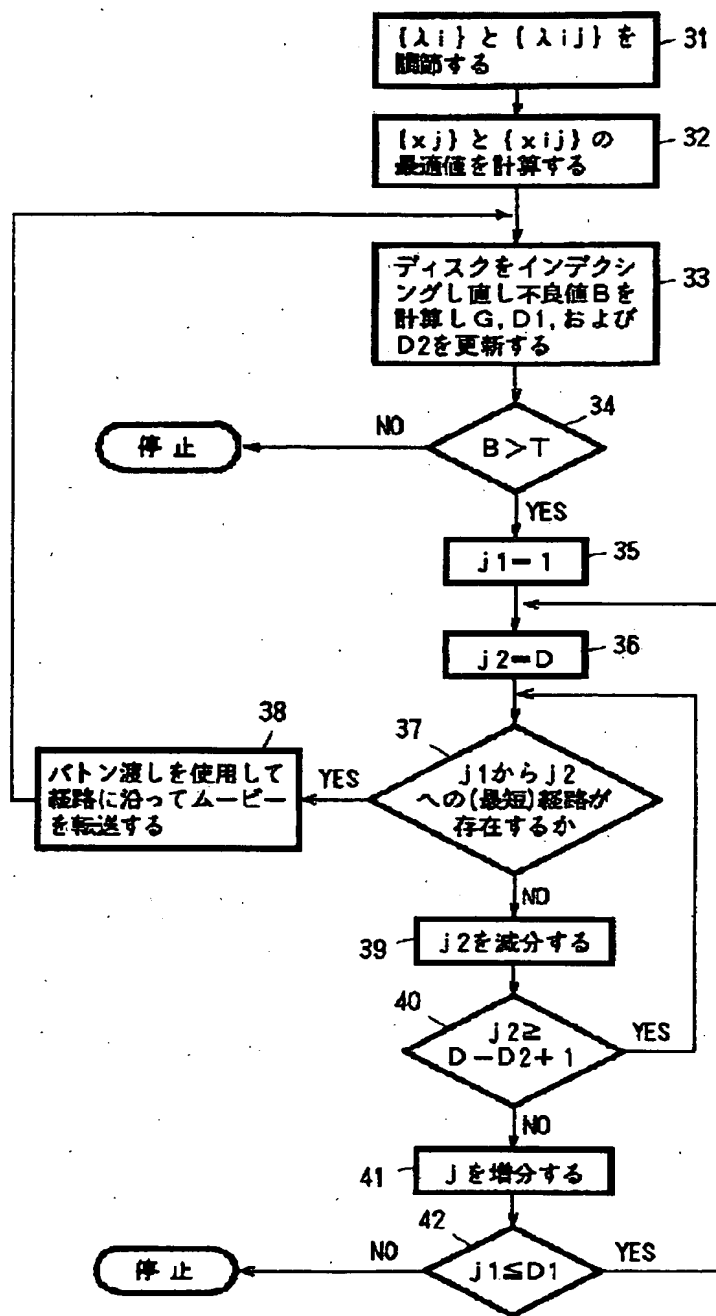
【 図4 】



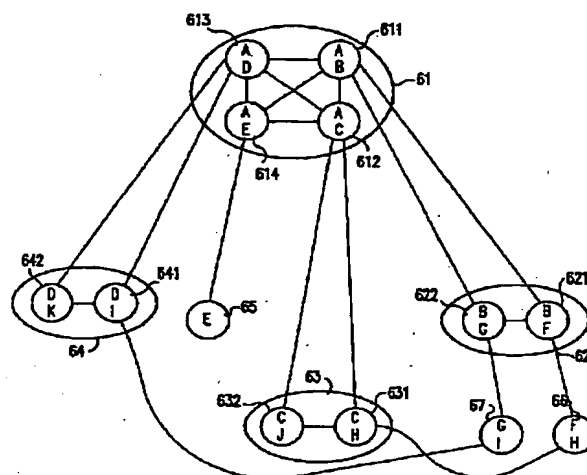
【 図2 】



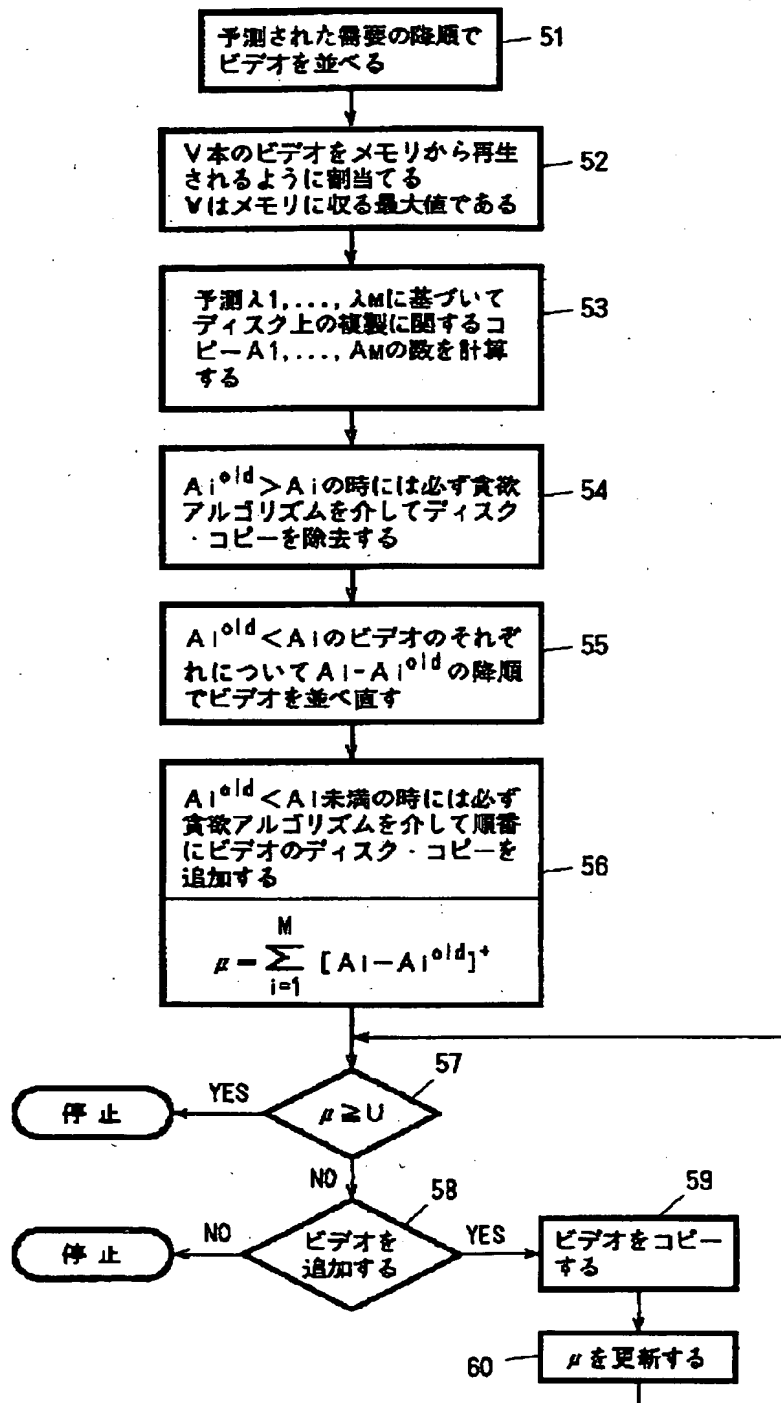
【 図3 】



【 図6 】



【 図5 】



フロント ページの続き

(72)発明者 ジョエル・レオナード・ウルフ  
アメリカ合衆国10526 ニューヨーク州ゴ  
ールデンズ・ブリッジ チェロキー・コー  
ト 7

(72)発明者 フィリップ・シーニルン・ユー  
アメリカ合衆国10514 ニューヨーク州チ  
ャパクア ストローノウェイ 18



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ **BLACK BORDERS**

☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☒ **FADED TEXT OR DRAWING**

☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**